

SEQUENCE LISTING

<110> Thorgeirsson, Snorri S.
 Woitach, Joseph T.
 Zhang, Minghuang

<120> cDNA ENCODING A GENE BOG (B5T OVER-EXPRESSED GENE) AND ITS PROTEIN
 PRODUCT

<130> 11613.29USW1

<140> US 09/637,746
 <141> 2000-08-11

<150> PCT/US99/04142
 <151> 1999-02-25

<150> US 60/079,567
 <151> 1998-03-27

<150> US 60/075,922
 <151> 1998-02-25

<160> 15

<170> PatentIn version 3.1

<210> 1
 <211> 1897
 <212> DNA
 <213> Rattus norvegicus

<220>
 <221> CDS
 <222> (18)..(536)
 <223>

<400> 1
 gtctcacgctc tgcatta atg gtg tcc cct acc aag gca gtg att gtt cct 50
 Met Val Ser Pro Thr Lys Ala Val Ile Val Pro
 1 5 10

ggg aac gga ggc ggg gat gtg gcc acc cac ggc tgg tac ggc tgg gtg 98
 Gly Asn Gly Gly Gly Asp Val Ala Thr His Gly Trp Tyr Gly Trp Val
 15 20 25

aga aag ggg ctg gag cag att cct ggt ttc cag tgt ttg gct aaa aac 146
 Arg Lys Gly Leu Glu Gln Ile Pro Gly Phe Gln Cys Leu Ala Lys Asn
 30 35 40

atg cct gac cca att acc gct cga gag agc atc tgg ctg ccc ttc atg 194
 Met Pro Asp Pro Ile Thr Ala Arg Glu Ser Ile Trp Leu Pro Phe Met
 45 50 55

gag aca gaa ctg cac tgt gat gag aag acc atc atc ata ggc cac agt 242

Glu Thr Glu Leu His Cys Asp Glu Lys Thr Ile Ile Ile Gly His Ser	
60 65 70 75	
tcc ggg gcc atc gca gcc atg agg tat gca gag aca cat cag gta tac	290
Ser Gly Ala Ile Ala Ala Met Arg Tyr Ala Glu Thr His Gln Val Tyr	
80 85 90	
gct ctc ata ttg gtg tct gca tac aca tca gac ttg gga gat gaa aat	338
Ala Leu Ile Leu Val Ser Ala Tyr Thr Ser Asp Leu Gly Asp Glu Asn	
95 100 105	
gag cgt gca agt ggg tac ttc agc cgc ccc tgg cag tgg gag aag atc	386
Glu Arg Ala Ser Gly Tyr Phe Ser Arg Pro Trp Gln Trp Glu Lys Ile	
110 115 120	
aag gcc aac tgc cct cac att ata cag ttt ggc tct act gat gac ccc	434
Lys Ala Asn Cys Pro His Ile Ile Gln Phe Gly Ser Thr Asp Asp Pro	
125 130 135	
ttc ctt cca tgg aag gaa caa caa gaa gtg gca gat agc tgg acg cca	482
Phe Leu Pro Trp Lys Glu Gln Gln Glu Val Ala Asp Ser Trp Thr Pro	
140 145 150 155	
aac tgt aca aat tca ctg acc gtg gtc act ttc aga aca cag agt tcc	530
Asn Cys Thr Asn Ser Leu Thr Val Val Thr Phe Arg Thr Gln Ser Ser	
160 165 170	
atg aac tgattagagt ggtgaagtct atgctgactc ctgctctgta acacgccagg	586
Met Asn	
atggggtaga agagtaacag ccgctaccct cacacagctt agacatggac gtccgtccag	646
ttagactaca gaagtgtctg agcaacaaac ccatttgaac actcacactg agttagtagc	706
acttccagtt cccacagagc ttaaattctcc ccaaaagcta ctagctacag cagtatgttt	766
cctgtttgat aagagacagg ttttttatatt ttaagctatc ctgttgatgc aaagagagtt	826
aagtcagaag aatccagaac ttgacataga cctggttgtg tgtccctgta atcatttcag	886
aaagcagggt caggagggaa ggctatcttg accctgtctc agaaagagtg agcaagaaga	946
tgaccaggt ctctgaggc tcattccaaa ttataactca ctatgttttag caagatgtgt	1006
tccacttctg agaccccagt acttgaaaa ctgaggaagg ttcattctga gtttgagacc	1066
ttgtctaggc taagtaaacc ctgtctcaaa acaacaacaa aaacaggttt tcattcaact	1126
tatatgactg acactttcca tttgtaataa aaaattttct cttactgggg aaatgaaaac	1186
acgattcaag gtccagaatg ttgtcttaga actcaaaact ctggttgctc tttaaaactg	1246
gctcaagaga ataaactcaa acttgggtgt tcatcattga aattcctgac cccaccacgt	1306
ccccaaccgt ccagactcta cagtgagagt gacacatatg atgctagtag actgcaggca	1366

gtatctgtta tacactgtat agactgcaga ttcgatcatg ggagtgtctgc aatataaaaa 1426
 tgtgacctat gtctttttta ctagagaata tagtgtgtat ataattccta catgaattat 1486
 ggtaactggg aacagcattg taattaaaag atttgcaa at gctactacaa gacaaagacc 1546
 aggtcatccc tttgtgaact tggtcgtaaa catttttaga atctctatga agtccaagaa 1606
 aaacaagata actaaaatga cataatacta aagggtggaa aacaaggagc aatcgtat 1666
 tgttattaag tttttaagta tcttcaaaag aacttttcca gggctaggga gaaggctgac 1726
 agtcaagagg ctactgagtt tttttccaga gttctgagtt caattcccag caactacatg 1786
 gtagtcacaa ccatctgtaa tggaccogat gccctcttct ggtgtgtctg aagacagcta 1846
 cagtgacttc acatacataa aataagtaaa tcttaaaaaa aaaaaaaaaa a 1897

<210> 2
 <211> 173
 <212> PRT
 <213> Rattus norvegicus

<400> 2

Met Val Ser Pro Thr Lys Ala Val Ile Val Pro Gly Asn Gly Gly Gly
 1 5 10 15

Asp Val Ala Thr His Gly Trp Tyr Gly Trp Val Arg Lys Gly Leu Glu
 20 25 30

Gln Ile Pro Gly Phe Gln Cys Leu Ala Lys Asn Met Pro Asp Pro Ile
 35 40 45

Thr Ala Arg Glu Ser Ile Trp Leu Pro Phe Met Glu Thr Glu Leu His
 50 55 60

Cys Asp Glu Lys Thr Ile Ile Ile Gly His Ser Ser Gly Ala Ile Ala
 65 70 75 80

Ala Met Arg Tyr Ala Glu Thr His Gln Val Tyr Ala Leu Ile Leu Val
 85 90 95

Ser Ala Tyr Thr Ser Asp Leu Gly Asp Glu Asn Glu Arg Ala Ser Gly
 100 105 110

Tyr Phe Ser Arg Pro Trp Gln Trp Glu Lys Ile Lys Ala Asn Cys Pro
115 120 125

His Ile Ile Gln Phe Gly Ser Thr Asp Asp Pro Phe Leu Pro Trp Lys
130 135 140

Glu Gln Gln Glu Val Ala Asp Ser Trp Thr Pro Asn Cys Thr Asn Ser
145 150 155 160

Leu Thr Val Val Thr Phe Arg Thr Gln Ser Ser Met Asn
165 170

<210> 3
<211> 98
<212> PRT
<213> Human papillomavirus

<400> 3

Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu Gln
1 5 10 15

Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser Ser
20 25 30

Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro Asp
35 40 45

Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp Ser Thr
50 55 60

Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu Glu
65 70 75 80

Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser Gln
85 90 95

Lys Pro

<210> 4
<211> 20
<212> PRT
<213> Simian virus

<400> 4

Asn Ala Phe Asn Glu Glu Asn Leu Phe Cys Ser Glu Glu Met Pro Ser
1 5 10 15

Ser Asp Asp Glu
20

<210> 5
<211> 22
<212> PRT
<213> Adenovirus

<400> 5

Asn Leu Val Pro Glu Val Ile Asp Leu Thr Cys His Glu Ala Gly Phe
1 5 10 15

Pro Pro Ser Asp Asp Glu
20

<210> 6
<211> 19
<212> PRT
<213> Homo sapiens

<400> 6

Leu Ile Gly Pro Glu Thr Leu Val Cys His Glu Val Asp Leu Thr Ser
1 5 10 15

Glu Ile Asp

<210> 7
<211> 522
<212> DNA
<213> Homo sapiens

<400> 7
atggtgtccc ccagcaaggc agtgattggt cccgggaaga taggtgggga tgagaccacc 60
cacggctggt atggctgggt gaaaaaggag ctggagaaga tacctggttt ccagtgtttg 120
gctaaaaaca tgcccgaccc aattaccgcg cgagagagca tctggctgcc cttcatggag 180
acagaactgc actgtgatga gaagactatc atcattggcc acagttccgg ggccatcgcg 240
gccatgaggt atgcagaaac acatcgagta tatgctctca tattggtgtc tgcatacaca 300

tcagagtttg gagatgaaaa tgagcgtgca agtgggtact tcagccgccc ctggcagtgg 360
gagaagatca aggccaaactg ccctcacatt gtacagtttg gctctactga tgacccttc 420
cttccttgga aggaacaaca agaagtggca gatagctgga cgccaaattg taaaaattca 480
ctgaccgtgg tcactttcag aacacagagt tccatgaact ga 522

<210> 8
<211> 173
<212> PRT
<213> Homo sapiens

<400> 8

Met Val Ser Pro Ser Lys Ala Val Ile Val Pro Gly Lys Ile Gly Gly
1 5 10 15

Asp Glu Thr Thr His Gly Trp Tyr Gly Trp Val Lys Lys Glu Leu Glu
20 25 30

Lys Ile Pro Gly Phe Gln Cys Leu Ala Lys Asn Met Pro Asp Pro Ile
35 40 45

Thr Ala Arg Glu Ser Ile Trp Leu Pro Phe Met Glu Thr Glu Leu His
50 55 60

Cys Asp Glu Lys Thr Ile Ile Ile Gly His Ser Ser Gly Ala Ile Ala
65 70 75 80

Ala Met Arg Tyr Ala Glu Thr His Arg Val Tyr Ala Leu Ile Leu Val
85 90 95

Ser Ala Tyr Thr Ser Glu Phe Gly Asp Glu Asn Glu Arg Ala Ser Gly
100 105 110

Tyr Phe Ser Arg Pro Trp Gln Trp Glu Lys Ile Lys Ala Asn Cys Pro
115 120 125

His Ile Val Gln Phe Gly Ser Thr Asp Asp Pro Phe Leu Pro Trp Lys
130 135 140

Glu Gln Gln Glu Val Ala Asp Ser Trp Thr Pro Asn Cys Thr Asn Ser
145 150 155 160

Leu Thr Val Val Thr Phe Arg Thr Gln Ser Ser Met Asn
165 170

<210> 9
<211> 522
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (1)..(522)
<223>

<400> 9
atg gcg tcc ccc aac aag gca gtg att gtt cct ggg aac gga ggc ggg 48
Met Ala Ser Pro Asn Lys Ala Val Ile Val Pro Gly Asn Gly Gly Gly
1 5 10 15
gat gtg gcc acc cac ggc tgg tat ggc tgg gtg aaa aag ggg ctg gag 96
Asp Val Ala Thr His Gly Trp Tyr Gly Trp Val Lys Lys Gly Leu Glu
20 25 30
cag att cct ggt ttc cag tgt ttg gct aaa aac atg cct gac cca att 144
Gln Ile Pro Gly Phe Gln Cys Leu Ala Lys Asn Met Pro Asp Pro Ile
35 40 45
acc gcg cga gag agc atc tgg ctg ccc ttc atg gag aca gag ctg cac 192
Thr Ala Arg Glu Ser Ile Trp Leu Pro Phe Met Glu Thr Glu Leu His
50 55 60
tgt gac gag aag acc atc atc ata ggc cac agt tcc ggg gcc atc gca 240
Cys Asp Glu Lys Thr Ile Ile Ile Gly His Ser Ser Gly Ala Ile Ala
65 70 75 80
gcc atg agg tat gca gag aca cat cag gta tac gct ctc gta ttg gtg 288
Ala Met Arg Tyr Ala Glu Thr His Gln Val Tyr Ala Leu Val Leu Val
85 90 95
tct gca tac aca tca gac ttg gga gat gaa aat gag cgt gca agt ggg 336
Ser Ala Tyr Thr Ser Asp Leu Gly Asp Glu Asn Glu Arg Ala Ser Gly
100 105 110
tac ttc agc cgc ccc tgg cag tgg gag aag atc aag gcc aac tgc cct 384
Tyr Phe Ser Arg Pro Trp Gln Trp Glu Lys Ile Lys Ala Asn Cys Pro
115 120 125
cac att ata cag ttt ggc tct act gat gac ccc ttc ctt ccc tgg aag 432
His Ile Ile Gln Phe Gly Ser Thr Asp Asp Pro Phe Leu Pro Trp Lys
130 135 140
gaa caa caa gaa gtg gca gat agc tgg acg cca aat tgt aca aat tca 480
Glu Gln Gln Glu Val Ala Asp Ser Trp Thr Pro Asn Cys Thr Asn Ser
145 150 155 160

ctg acc gtg gtc act ttc aga aca cag agt tcc atg aac tga
 Leu Thr Val Val Thr Phe Arg Thr Gln Ser Ser Met Asn
 165 170

522

<210> 10
 <211> 173
 <212> PRT
 <213> Mus musculus

<400> 10

Met Ala Ser Pro Asn Lys Ala Val Ile Val Pro Gly Asn Gly Gly Gly
 1 5 10 15

Asp Val Ala Thr His Gly Trp Tyr Gly Trp Val Lys Lys Gly Leu Glu
 20 25 30

Gln Ile Pro Gly Phe Gln Cys Leu Ala Lys Asn Met Pro Asp Pro Ile
 35 40 45

Thr Ala Arg Glu Ser Ile Trp Leu Pro Phe Met Glu Thr Glu Leu His
 50 55 60

Cys Asp Glu Lys Thr Ile Ile Ile Gly His Ser Ser Gly Ala Ile Ala
 65 70 75 80

Ala Met Arg Tyr Ala Glu Thr His Gln Val Tyr Ala Leu Val Leu Val
 85 90 95

Ser Ala Tyr Thr Ser Asp Leu Gly Asp Glu Asn Glu Arg Ala Ser Gly
 100 105 110

Tyr Phe Ser Arg Pro Trp Gln Trp Glu Lys Ile Lys Ala Asn Cys Pro
 115 120 125

His Ile Ile Gln Phe Gly Ser Thr Asp Asp Pro Phe Leu Pro Trp Lys
 130 135 140

Glu Gln Gln Glu Val Ala Asp Ser Trp Thr Pro Asn Cys Thr Asn Ser
 145 150 155 160

Leu Thr Val Val Thr Phe Arg Thr Gln Ser Ser Met Asn
 165 170

<210> 11
 <211> 714
 <212> DNA
 <213> Mus musculus

<400> 11
 cagagccctg aaaggttggt gcatgagccc gtgaaagtgg agtttcagtg gtagtggata 60
 gcataggaca ctggagacac agttcatgtc cagcattcat ggagtgggag cagagagttc 120
 cctgaagctc actggctagt attcttgcta aaccaatgag ctccaaattc acagatcttg 180
 tcgcaaaacc caaatgtaat gtggaaatga aggaaaagaa gacacccaac actgactgaa 240
 tatggtgaca ctccctttta atgccagcac tcaggagaca aaaagcaggc agatcttttg 300
 tgagttctag gccagtctgg ttacataga cagctccagg ccagtaaggg gctacgtaat 360
 gaaactgtct taaacaaatt aaggaacgtt catttgaaaa aaaataaacc ttccttaaag 420
 aagtattggt acaactaata aaaagataac acattatgag cacgctgttg ccagcacata 480
 agggatgtgg agtatgagaa cgctggaaaa ggggtaaatc aaagataatt aatatttgat 540
 ggtaattcac aggtttgagt ttagctgcct gtgcttttagc cagaaaatgc gtaggcctgc 600
 aggtatccaa gaactacaat tcccagaagt ccgcagtgca ggctctgggc cggatgtagt 660
 cttggtctga gagctgctgg tccaagctgg gcaaggtctc ccacgtctac attc 714

<210> 12
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Antisense

<400> 12
 aatcactgcc ttggtagggg acaccattaa 30

<210> 13
 <211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Antisense

<400> 13
 tcagttcatg gaactctgtg ttctgaaagt gac 33

<210> 14
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 14
atggtctctc ctagc 15

<210> 15
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 15
gagttccatg aac 13

10